

SEQUENCE LISTING

<110> Conklin, Darrell C.
Gao, Zeren

<120> MULTI-DOMAIN PROTEINASE INHIBITOR

<130> 00-25

<150> US 60/193,642

<151> 2000-03-31

<160> 13

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 2082

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (376)...(2022)

<400> 1

```

gtgaccctca tggccagtgg ctctgtgctc atgggcctct ggccctccc caacctcctc   60
ccctctgccc tgtctgacc agggcctggg agccccgca cggttcagac agaggggcca   120
ggctgaagct ggagaggaac cagcgtcaca cagacggcct ctgagaactt ggagaccccg   180
ttaccacccc agcaggggtg tcaggacaag catctgctgc aggcttcagc ctcaggggca   240
aaagggagcc ccggggctct ggtgggggca ccgaccacag gcccgagggt tggatgcctg   300
caggaagctg ggctctgtgg agcccgagga ggggctggtg gccacacccc ccggccccct   360
ggctcggcgg ccctc atg ccc gcc cta cgt cca ctc ctg ccg ctc ttg ctc   411
                Met Pro Ala Leu Arg Pro Leu Leu Pro Leu Leu Leu
                  1             5                 10

ctc ctc cgg ctg acc tcg ggg gct ggc ttg ctg cca ggg ctg ggg agc   459
Leu Leu Arg Leu Thr Ser Gly Ala Gly Leu Leu Pro Gly Leu Gly Ser
    15                 20                 25

```

cac ccg ggc gtg tgc ccc aac cag ctc agc ccc aac ctg tgg gtg gac	507
His Pro Gly Val Cys Pro Asn Gln Leu Ser Pro Asn Leu Trp Val Asp	
30 35 40	
gcc cag agc acc tgt gag cgc gag tgt agc agg gac cag gac tgt gcg	555
Ala Gln Ser Thr Cys Glu Arg Glu Cys Ser Arg Asp Gln Asp Cys Ala	
45 50 55 60	
gct gct gag aag tgc tgc atc aac gtg tgt gga ctg cac agc tgc gtg	603
Ala Ala Glu Lys Cys Cys Ile Asn Val Cys Gly Leu His Ser Cys Val	
65 70 75	
gca gca cgc ttc ccc ggc agc cca gct gcg ccg acg aca gcg gcc tcc	651
Ala Ala Arg Phe Pro Gly Ser Pro Ala Ala Pro Thr Thr Ala Ala Ser	
80 85 90	
tgc gag ggc ttt gtg tgc cca cag cag ggc tcg gac tgc gac atc tgg	699
Cys Glu Gly Phe Val Cys Pro Gln Gln Gly Ser Asp Cys Asp Ile Trp	
95 100 105	
gac ggg cag ccc gtg tgc cgc tgc cgc gac cgc tgt gag aag gag ccc	747
Asp Gly Gln Pro Val Cys Arg Cys Arg Asp Arg Cys Glu Lys Glu Pro	
110 115 120	
agc ttc acc tgc gcc tcg gac ggc ctc acc tac tac aac cgc tgc tat	795
Ser Phe Thr Cys Ala Ser Asp Gly Leu Thr Tyr Tyr Asn Arg Cys Tyr	
125 130 135 140	
atg gac gcc gag gcc tgc ctg cgg ggc ctg cac ctc cac atc gtg ccc	843
Met Asp Ala Glu Ala Cys Leu Arg Gly Leu His Leu His Ile Val Pro	
145 150 155	
tgc aag cac gtg ctc agc tgg ccg ccc agc agc ccg ggg ccg ccg gag	891
Cys Lys His Val Leu Ser Trp Pro Pro Ser Ser Pro Gly Pro Pro Glu	
160 165 170	
acc act gcc cgc ccc aca cct ggg gcc gcg ccc gtg cct cct gcc ctg	939
Thr Thr Ala Arg Pro Thr Pro Gly Ala Ala Pro Val Pro Pro Ala Leu	
175 180 185	
tac agc agc ccc tcc cca cag gcg gtg cag gtt ggg ggt acg gcc agc	987
Tyr Ser Ser Pro Ser Pro Gln Ala Val Gln Val Gly Gly Thr Ala Ser	
190 195 200	

ctc cac tgc gac gtc agc ggc cgc ccg ccg cct gct gtg acc tgg gag	1035
Leu His Cys Asp Val Ser Gly Arg Pro Pro Pro Ala Val Thr Trp Glu	
205 210 215 220	
aag cag agt cac cag cga gag aac ctg atc atg cgc cct gat cag atg	1083
Lys Gln Ser His Gln Arg Glu Asn Leu Ile Met Arg Pro Asp Gln Met	
225 230 235	
tat ggc aac gtg gtg gtc acc agc atc ggg cag ctg gtg ctc tac aac	1131
Tyr Gly Asn Val Val Val Thr Ser Ile Gly Gln Leu Val Leu Tyr Asn	
240 245 250	
gcg cgg ccc gaa gac gcc ggc ctg tac acc tgc acc gcg cgc aac gct	1179
Ala Arg Pro Glu Asp Ala Gly Leu Tyr Thr Cys Thr Ala Arg Asn Ala	
255 260 265	
gct ggg ctg ctg cgg gct gac ttc cca ctc tct gtg gtc cag cga gag	1227
Ala Gly Leu Leu Arg Ala Asp Phe Pro Leu Ser Val Val Gln Arg Glu	
270 275 280	
ccg gcc agg gac gca gcc ccc agc atc cca gcc ccg gcc gag tgc ctg	1275
Pro Ala Arg Asp Ala Ala Pro Ser Ile Pro Ala Pro Ala Glu Cys Leu	
285 290 295 300	
ccg gat gtg cag gcc tgc acg ggc ccc act tcc cca cac ctt gtc ctc	1323
Pro Asp Val Gln Ala Cys Thr Gly Pro Thr Ser Pro His Leu Val Leu	
305 310 315	
tgg cac tac gac ccg cag cgg ggc ggc tgc atg acc ttc ccg gcc cgt	1371
Trp His Tyr Asp Pro Gln Arg Gly Gly Cys Met Thr Phe Pro Ala Arg	
320 325 330	
ggc tgt gat ggg gcg gcc cgc ggc ttt gag acc tac gag gca tgc cag	1419
Gly Cys Asp Gly Ala Ala Arg Gly Phe Glu Thr Tyr Glu Ala Cys Gln	
335 340 345	
cag gcc tgt gcc cgc ggc ccc ggc gac gcc tgc gtg ctg cct gcc gtg	1467
Gln Ala Cys Ala Arg Gly Pro Gly Asp Ala Cys Val Leu Pro Ala Val	
350 355 360	

cag ggc ccc tgc cgg ggc tgg gag ccg cgc tgg gcc tac agc ccg ctg Gln Gly Pro Cys Arg Gly Trp Glu Pro Arg Trp Ala Tyr Ser Pro Leu 365 370 375 380	1515
ctg cag cag tgc cat ccc ttc gtg tac ggt ggc tgc gag ggc aac ggc Leu Gln Gln Cys His Pro Phe Val Tyr Gly Gly Cys Glu Gly Asn Gly 385 390 395	1563
aac aac ttc cac agc cgc gag agc tgc gag gat gcc tgc ccc gtg ccg Asn Asn Phe His Ser Arg Glu Ser Cys Glu Asp Ala Cys Pro Val Pro 400 405 410	1611
cgc aca ccg ccc tgc cgc gcc tgc cgc ctc cgg agc aag ctg gcg ctg Arg Thr Pro Pro Cys Arg Ala Cys Arg Leu Arg Ser Lys Leu Ala Leu 415 420 425	1659
agc ctg tgc cgc agc gac ttc gcc atc gtg ggg cgg ctc acg gag gtg Ser Leu Cys Arg Ser Asp Phe Ala Ile Val Gly Arg Leu Thr Glu Val 430 435 440	1707
ctg gag gag ccc gag gcc gcc ggc ggc atc gcc cgc gtg gcg ctc gag Leu Glu Glu Pro Glu Ala Ala Gly Gly Ile Ala Arg Val Ala Leu Glu 445 450 455 460	1755
gac gtg ctc aag gat gac aag atg ggc ctc aag ttc ttg ggc acc aag Asp Val Leu Lys Asp Asp Lys Met Gly Leu Lys Phe Leu Gly Thr Lys 465 470 475	1803
tac ctg gag gtg acg ctg agt ggc atg gac tgg gcc tgc ccc tgc ccc Tyr Leu Glu Val Thr Leu Ser Gly Met Asp Trp Ala Cys Pro Cys Pro 480 485 490	1851
aac atg acg gcg ggc gac ggg ccg ctg gtc atc atg ggt gag gtg cgc Asn Met Thr Ala Gly Asp Gly Pro Leu Val Ile Met Gly Glu Val Arg 495 500 505	1899
gat ggc gtg gcc gtg ctg gac gcc ggc agc tac gtc cgc gcc gcc agc Asp Gly Val Ala Val Leu Asp Ala Gly Ser Tyr Val Arg Ala Ala Ser 510 515 520	1947
gag aag cgc gtc aag aag atc ttg gag ctg ctg gag aag cag gcc tgc Glu Lys Arg Val Lys Lys Ile Leu Glu Leu Leu Glu Lys Gln Ala Cys 525 530 535 540	1995

gag ctg ctc aac cgc ttc cag gac tag ccccgagagg ggcctgcgcc 2042
 Glu Leu Leu Asn Arg Phe Gln Asp *

545

accccgctcct ggtgaataaa cgcactccct gtgcctcaga 2082

<210> 2

<211> 548

<212> PRT

<213> Homo sapiens

<400> 2

Met Pro Ala Leu Arg Pro Leu Leu Pro Leu Leu Leu Leu Arg Leu
 1 5 10 15
 Thr Ser Gly Ala Gly Leu Leu Pro Gly Leu Gly Ser His Pro Gly Val
 20 25 30
 Cys Pro Asn Gln Leu Ser Pro Asn Leu Trp Val Asp Ala Gln Ser Thr
 35 40 45
 Cys Glu Arg Glu Cys Ser Arg Asp Gln Asp Cys Ala Ala Ala Glu Lys
 50 55 60
 Cys Cys Ile Asn Val Cys Gly Leu His Ser Cys Val Ala Ala Arg Phe
 65 70 75 80
 Pro Gly Ser Pro Ala Ala Pro Thr Thr Ala Ala Ser Cys Glu Gly Phe
 85 90 95
 Val Cys Pro Gln Gln Gly Ser Asp Cys Asp Ile Trp Asp Gly Gln Pro
 100 105 110
 Val Cys Arg Cys Arg Asp Arg Cys Glu Lys Glu Pro Ser Phe Thr Cys
 115 120 125
 Ala Ser Asp Gly Leu Thr Tyr Tyr Asn Arg Cys Tyr Met Asp Ala Glu
 130 135 140
 Ala Cys Leu Arg Gly Leu His Leu His Ile Val Pro Cys Lys His Val
 145 150 155 160
 Leu Ser Trp Pro Pro Ser Ser Pro Gly Pro Pro Glu Thr Thr Ala Arg
 165 170 175
 Pro Thr Pro Gly Ala Ala Pro Val Pro Pro Ala Leu Tyr Ser Ser Pro
 180 185 190
 Ser Pro Gln Ala Val Gln Val Gly Gly Thr Ala Ser Leu His Cys Asp
 195 200 205
 Val Ser Gly Arg Pro Pro Pro Ala Val Thr Trp Glu Lys Gln Ser His
 210 215 220
 Gln Arg Glu Asn Leu Ile Met Arg Pro Asp Gln Met Tyr Gly Asn Val
 225 230 235 240

Val Val Thr Ser Ile Gly Gln Leu Val Leu Tyr Asn Ala Arg Pro Glu
 245 250 255
 Asp Ala Gly Leu Tyr Thr Cys Thr Ala Arg Asn Ala Ala Gly Leu Leu
 260 265 270
 Arg Ala Asp Phe Pro Leu Ser Val Val Gln Arg Glu Pro Ala Arg Asp
 275 280 285
 Ala Ala Pro Ser Ile Pro Ala Pro Ala Glu Cys Leu Pro Asp Val Gln
 290 295 300
 Ala Cys Thr Gly Pro Thr Ser Pro His Leu Val Leu Trp His Tyr Asp
 305 310 315 320
 Pro Gln Arg Gly Gly Cys Met Thr Phe Pro Ala Arg Gly Cys Asp Gly
 325 330 335
 Ala Ala Arg Gly Phe Glu Thr Tyr Glu Ala Cys Gln Gln Ala Cys Ala
 340 345 350
 Arg Gly Pro Gly Asp Ala Cys Val Leu Pro Ala Val Gln Gly Pro Cys
 355 360 365
 Arg Gly Trp Glu Pro Arg Trp Ala Tyr Ser Pro Leu Leu Gln Gln Cys
 370 375 380
 His Pro Phe Val Tyr Gly Gly Cys Glu Gly Asn Gly Asn Asn Phe His
 385 390 395 400
 Ser Arg Glu Ser Cys Glu Asp Ala Cys Pro Val Pro Arg Thr Pro Pro
 405 410 415
 Cys Arg Ala Cys Arg Leu Arg Ser Lys Leu Ala Leu Ser Leu Cys Arg
 420 425 430
 Ser Asp Phe Ala Ile Val Gly Arg Leu Thr Glu Val Leu Glu Glu Pro
 435 440 445
 Glu Ala Ala Gly Gly Ile Ala Arg Val Ala Leu Glu Asp Val Leu Lys
 450 455 460
 Asp Asp Lys Met Gly Leu Lys Phe Leu Gly Thr Lys Tyr Leu Glu Val
 465 470 475 480
 Thr Leu Ser Gly Met Asp Trp Ala Cys Pro Cys Pro Asn Met Thr Ala
 485 490 495
 Gly Asp Gly Pro Leu Val Ile Met Gly Glu Val Arg Asp Gly Val Ala
 500 505 510
 Val Leu Asp Ala Gly Ser Tyr Val Arg Ala Ala Ser Glu Lys Arg Val
 515 520 525
 Lys Lys Ile Leu Glu Leu Leu Glu Lys Gln Ala Cys Glu Leu Leu Asn
 530 535 540
 Arg Phe Gln Asp
 545

<210> 3

<211> 55

<212> PRT

<213> Homo sapiens

<400> 3

```

Thr Asp Ile Cys Lys Leu Pro Lys Asp Glu Gly Thr Cys Arg Asp Phe
 1           5           10           15
Ile Leu Lys Trp Tyr Tyr Asp Pro Asn Thr Lys Ser Cys Ala Arg Phe
          20           25           30
Trp Tyr Gly Gly Cys Gly Gly Asn Glu Asn Lys Phe Gly Ser Gln Lys
      35           40           45
Glu Cys Glu Lys Val Cys Ala
 50           55

```

<210> 4

<211> 6

<212> PRT

<213> Artificial Sequence

<220>

<223> Glu-Glu tag

<400> 4

```

Glu Tyr Met Pro Met Glu
 1           5

```

<210> 5

<211> 55

<212> PRT

<213> Artificial Sequence

<220>

<223> peptide motif

<221> VARIANT

<222> (2)...(7)

<223> Xaa = any amino acid

<221> VARIANT

<222> (8)...(9)

<223> Xaa = any amino acid or is not present

<221> VARIANT

<222> (11)...(25)

<223> Xaa = any amino acid

<221> VARIANT

<222> (26)...(29)

<223> Xaa = any amino acid or is not present

<221> VARIANT

<222> (31)...(54)

<223> Xaa = any amino acid

<400> 5

Cys	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Cys	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
1				5					10					15	
Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Cys	Xaa	Xaa
			20					25					30		
Xaa	Xaa	Xaa	Xaa	Xaa	Cys	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
			35				40						45		
Xaa	Xaa	Cys	Xaa	Xaa	Xaa	Cys									
			50				55								

<210> 6

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> thrombin cleavage site

<400> 6

Leu Val Pro Arg

1

<210> 7

<211> 1644

<212> DNA

<213> Artificial Sequence

<220>

<223> degenerate sequence

<221> misc_feature

<222> (1)...(1644)

<223> n = A,T,C or G

<400> 7

atgccngcny	tnmgncny	nytnccnytn	ytntytnytn	tnmgnytnac	nwsnggngcn	60
ggnytnytn	cnggnytnng	nwsncayccn	ggngtntgyc	cnaaycaryt	nwsnccnaay	120
ytntgggtng	aygcncarws	nacntgygar	mgngartgyw	snmgngayca	rgaytgygcn	180
gcngcngara	artgytgyat	haaygtntgy	ggnytncayw	sntgygtngc	ngcnmgntty	240
ccnggnwsnc	cngcngcncc	nacnacngcn	gcnwsntgyg	arggnttygt	ntgyccncar	300
carggnwsng	aytgygayat	htgggagyggn	carccngtnt	gymgntgymg	ngaymgntgy	360
garaargarc	cnwsnttyac	ntgygcwnsn	gayggnytna	cntaytayaa	ymgntgytay	420
atggaygcng	argcntgyyt	nmnggnytn	cayytncaya	thgtnccntg	yaarcaygtn	480
ytnwsntggc	cncnwsnws	nccnggnccn	ccngaracna	cngcnmgnc	nacnccnggn	540
gcngcnccng	tnccnccngc	nytnaywsn	wsnccnwsnc	cncargcngt	ncargtnggn	600
ggnacngcnw	snytncaytg	ygaygtwnsn	ggnmgnccnc	cncngcngt	nacntgggar	660
aarcarsnc	aycarmnga	raayytnath	atgmgnccng	aycaratgta	yggnaaygtn	720
gtngtnacnw	snathggnc	rytngtynytn	tayaaygcnm	gnccngarga	ygcnngnytn	780
tayacntgya	cngcnmgnaa	ygcnngnggn	ytntytnmgng	cngayttycc	nytnwsngtn	840
gtncarmng	arccngcnmg	ngaygcngcn	ccnwsnathc	cngcnccngc	ngartggytn	900
ccngaygtnc	argcntgyac	nggnccnacn	wsnccncayy	tngtynytn	gcaytaygay	960
ccncarmng	gnggntgyat	gacnttyccn	gcnmngngnt	gygayggngc	ngcnmgnggn	1020
ttygaracnt	aygargcntg	ycarcargcn	tgygcnmng	gnccnggnga	ygcntgygtn	1080
ytncngcng	tnccarggncc	ntgymngngn	tgggarccnm	gntgggcnta	ywsnccnytn	1140
ytncarcart	gycayccntt	ygntntaygg	ggntgygarg	gnaayggnaa	yaayttycay	1200
wsnmngarw	sntgygarga	ygcntgyccn	gtncnmgna	cncnccntg	ymngcngtgy	1260
mgnytnmgw	snaarytngc	nytnwsnytn	tgymgnwsng	ayttygcnat	hgtnggnmg	1320
ytnacngarg	tnytngarga	rccngargcn	gcngnggnga	thgcnmgnt	ngcnytngar	1380
gaygtnytna	argaygayaa	ratgggnytn	aarttyytn	gnacnaarta	yytnngargtn	1440
acnytnwsng	gnatggaytg	ggcntgyccn	tgyccnaaya	tgacngcngg	ngayggncn	1500
ytngtnatha	tggngarggt	nmngngaygg	gtngcngtny	tngaygcngg	nwsntaygtn	1560
mgngcngcnw	sngaraarmg	ngtnaaraar	athytngary	tnytngaraa	rcargcntgy	1620
garytnytna	aymgnttyca	rgay				1644

<210> 8

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> oligonucleotide primer ZC28,995

<400> 8

acttccccac accttgctct ct

22

<210> 9

<213> Artificial Sequence

<223> oligonucleotide primer ZC28,996

tgccctcgtag gtctcaaagc c

21

<213> Artificial Sequence

<223> oligonucleotide primer ZC29,898

gtcctctggc actacgaccc gca

23

<213> Artificial Sequence

<223> oligonucleotide primer ZC29,899

acggcaggca gcacgcagg

19

<213> Artificial Sequence

<223> oligonucleotide primer ZC37,161

cctgaccaaa tgtatggcaa cg

22

<211> 23

<212> DNA

<213> Artificial Sequence

 $\langle 220 \rangle$

<223> oligonucleotide primer ZC37,160

<400> 13

cctgggtccc tgcctgagt agt